A Color Stereo Image Set for Evaluation of Ordnance Recognition Algorithms

Clark F. Olson

Jet Propulsion Laboratory, California Institute of Technology

Mail Stop 107-102, 4800 Oak Grove Drive, Pasadena, CA 91109

http://robotics.jpl.nasa.gov/people/olson/homepage.html

February 24, 1998

Abstract

This report describes a dataset of color, stereo images of ordnance collected at a live-fire test range near Nellis Air Force Base. The images consist primarily of BLU-97 shells. 174 stereo pairs of images were collected, containing approximately 184 instances of BLU-97. The data was collected using a pair of Kodak DCS-410 digital cameras mounted rigidly on a tripod. Subsequent processing was performed to convert the images into a convenient format and to generate stereo range maps.

1 Introduction

Military test ranges containing unexploded ordnance due to live-fire testing and training exercises have become a significant safety problem in many locations. It has been estimated that over 800 locations and 11 million acres in the United States alone have been potentially contaminated with unexploded ordnance [2]. The cleanup of these sites has become of particular importance in this era of base realignment and closure. The estimated cleanup of sites in this category alone is over $4 billion. Furthermore, cleanup of such sites by human technicians is dangerous and the potential loss of human lives adds immeasurably to the cost.

We are in the process of developing an ordnance recognition system for use on an unmanned ground vehicle (UGV) in conjunction with the Rough Terrain Surface Munitions Clearance project at Wright Laboratory, Tyndall Air Force Base. The envisioned system consists of an unmanned all-terrain vehicle with on-board stereo cameras, in addition to verification and neutralization devices. The vehicle could either survey the test range, using the cameras as a pushbroom-like detection device, or move to a designated observation point and sweep the area of interest using pan and tilt.

Figure 1: Image of BLU-97 ordnance acquired at a Nellis Air Force Base test range.

We concentrate on the recognition of BLU-97 ordnance, which is in current usage in U.S. military test ranges (Figure 1). The body of this type of ordnance is cylindrical in shape, approximately 20 centimeters long, and has a 6 centimeter diameter. When new, the ordnance is bright yellow in color, but it is often weathered in practice on the test range.

In order for a system to be effective in this application, we estimate that the system should be able to recognize ordnance up to 10 meters from the vehicle and that a 5 meter field-of-view at this range is desirable. This implies that a lens with a 30° field-of-view or smaller should be used. In this case, a frontal view of the ordnance at the maximum range in a 500×500 image will yield a 20 pixel long strip in the image and non-frontal views will be even smaller.

We have begun algorithm development to solve this problem [5]. However, in order to evaluate the algorithms we must have some set of images upon which to perform testing. The report describes the data set that was collected so that this evaluation is possible. We have simulated the conditions for a UGV-mounted ordnance recognition system using a pair of digital cameras mounted on a tripod and collected a large set of imagery at a live-fire test range near Nellis Air Force Base.

The next section discusses the procedures used in collecting the data. Section 3 describes the data itself and gives annotations for the images. Section 4 discusses the stereo data that was extracted from the imagery, including the camera models used. Finally, Section 5 summarizes the report.

2 Procedures

Data was collected on the morning of 14 Oct 1997 at a test range near Nellis Air Force Base (a short distance from control point Bravo). The imagery was collected with a pair of Kodak DCS-410 digital cameras mounted rigidly on a tripod. The cameras had been previously calibrated to determine the intrinsic and extrinsic parameters (see Section 4).

Figure 2: Example of the terrain at the Air Force test range.

The terrain at this location consisted primarily of a dry mud-flat with many small cracks in the ground and with some debris. See Figure 2 for an example. Many instances of BLU-97 shells were present. A small number of BLU-38 bomblets[1] were also found.

The intent of the dataset was to cover as many variables as possible in the appearance of the ordnance, including ordnance position and orientation in the image, lighting direction, degree of ordnance discoloration, deformation, and obscuration. Variables that did not receive as much coverage as might be desired include the terrain (the images were obtained at a single location), the illumination (the images were obtained on a single cloudless day), and the sun angle (the images were acquired in the morning, before the sun was at full height, causing long shadows in some cases). Usually, several consecutive stereo pairs were collected of each piece of ordnance (or multiple closely aligned pieces of ordnance) with varying viewing distance and orientation. No precise ground-truth is available for the locations of the ordnance, although estimates of the image positions are given in the following section.

The cameras output the imagery in a complicated TIFF format in color at 1524×1012 pixels. For our purposes, we have transformed these images into a simple local format called CPIC, cut the effective size to 762×506 and padded the images on the right and bottom with blank space to a size of 768×512.

3 Data

The complete data set collected consists of 174 stereo pairs of color images, including two stereo pairs of a planar calibration target. Thumbnails of the complete set can be found in Figures 2-5. These images contain 184 instances of BLU-97 ordnance and 3 instances of BLU-38 ordnance. Of course, not all of the imagery taken is suitable for testing ordnance recognition. Not only are there several images of calibration targets and personnel, but

[1] These are baseball sized spheres (usually olive colored) with raised flanges.

a subset of the images on the test range have been discarded as not characteristic of a realistic scenario. Some reasons for discarding images include cases where the majority of the ordnance was obscured or outside of one the stereo images, cases where the operator misfired such that the stereo images were not in correspondence, and instances of ordnance that were severely self-shadowed. The culled data set consists of 123 images that were selected as representative of a realistic scenario and these images contain 119 BLU-97 instances. In addition, 10 of these images were selected as training images, which are to be used to optimize the parameters of the algorithms.

Tables 1-3 give annotations for the data set, including the type of ordnance present in the image, the rough location in the right image of the stereo pair, and notes describing the condition of the ordnance, the images that were culled from the set, and those selected as training images.

4 Stereo data

This section discusses the stereo data extracted from the imagery. First, we give a brief review of the camera model used and the parameters extracted from calibration.

4.1 Camera models

In order to generate stereo range maps, the cameras were calibrated using a camera model that allows arbitrary affine transformations of the image plane [6] and that has been extended to model radial lens distortion [1]. The linear model consists of four 3-vectors for each camera as follows:

- **c**: The focal center of the camera.
- **a**: The camera axis.
- **h**: A horizontal scale factor given by $\mathbf{h} = \mathbf{h}' + x_c\mathbf{a}$, where x_c is the horizontal center of the image, and $\mathbf{h}'$ is a vector parallel to the image x-axis with length given by the amount of change in the x-coordinate of a point $\mathbf{p}$ that is caused by a change of one in the tangent of the angle $\angle\mathbf{acp}$ as $\mathbf{p}$ moves horizontally in the image.
- **v**: A vertical scale factor given by $\mathbf{v} = \mathbf{v}' + x_c\mathbf{a}$, where y_c is the vertical center of the image, and $\mathbf{v}'$ is a vector parallel to the image y-axis with length given by the amount of change in the y-coordinate of a point $\mathbf{p}$ that is caused by a change of one in the tangent of the angle $\angle\mathbf{acp}$ as $\mathbf{p}$ moves vertically in the image.

With this formulation, the image coordinates of a point $\mathbf{p}$ in free space can be written conveniently as follows:

$$x = \frac{(\mathbf{p} - \mathbf{c}) \cdot \mathbf{h}}{(\mathbf{p} - \mathbf{c}) \cdot \mathbf{a}}$$

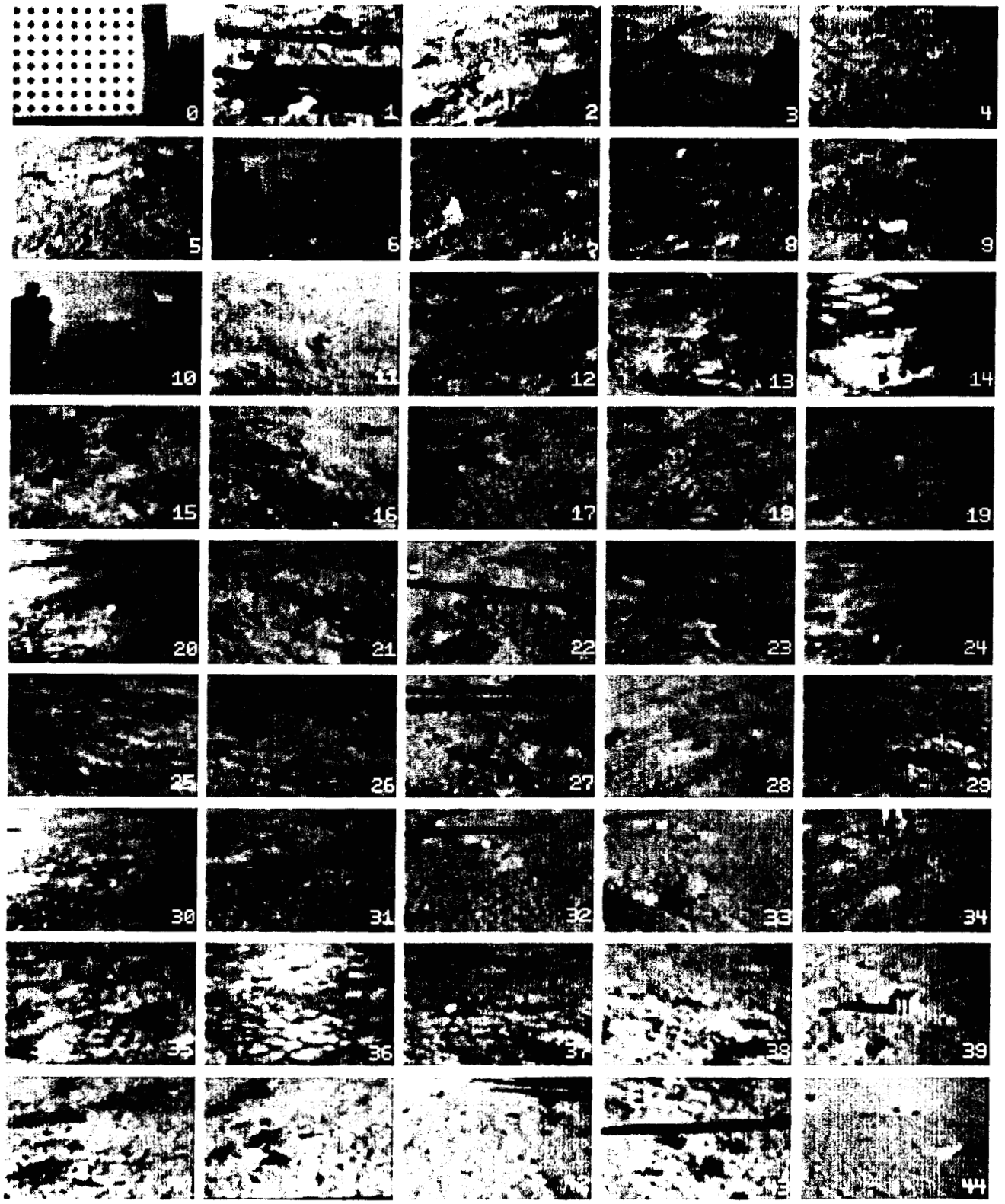

Figure 3: Data set thumbnails 0-44.

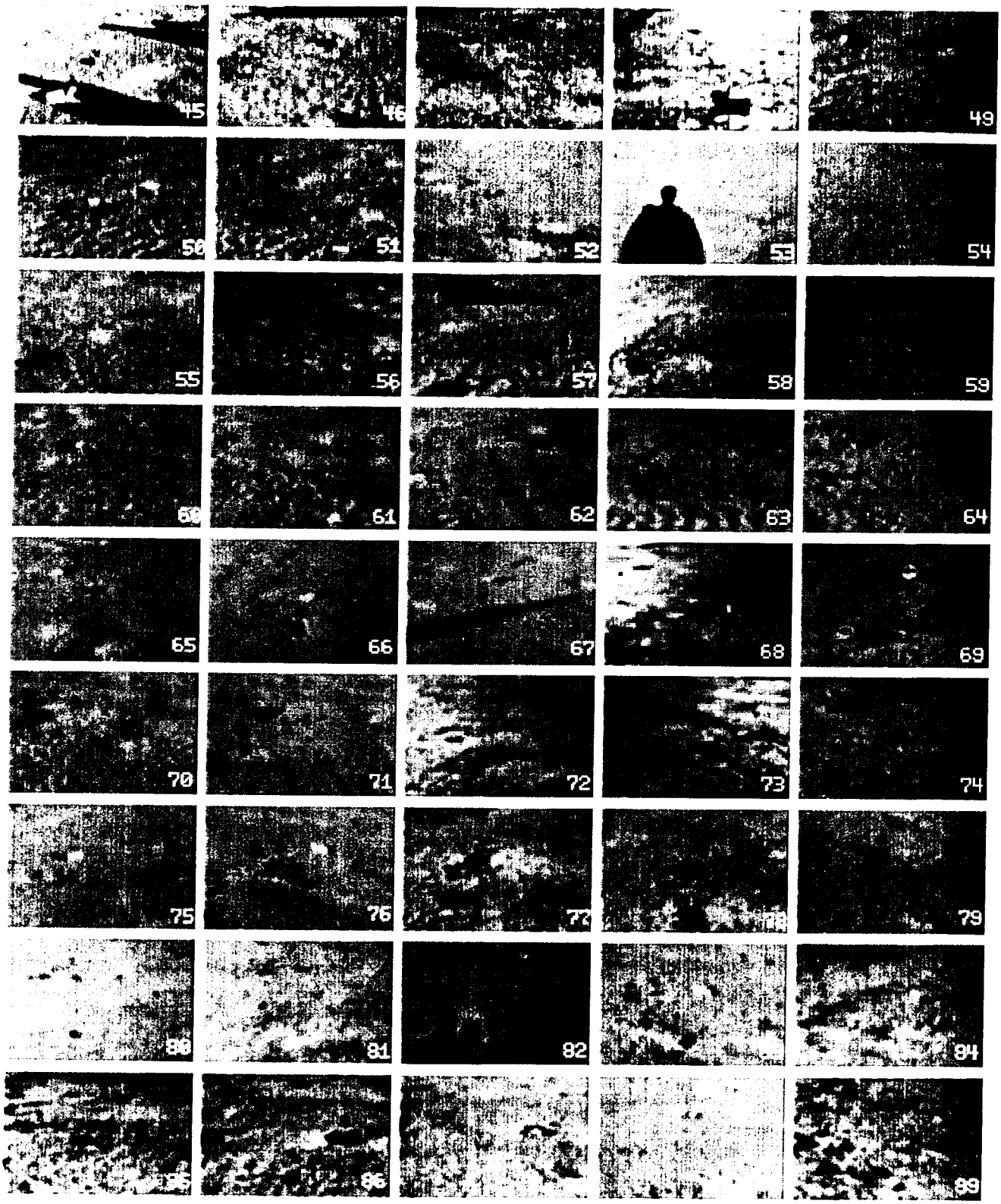

Figure 4: Data set thumbnails 45-89.

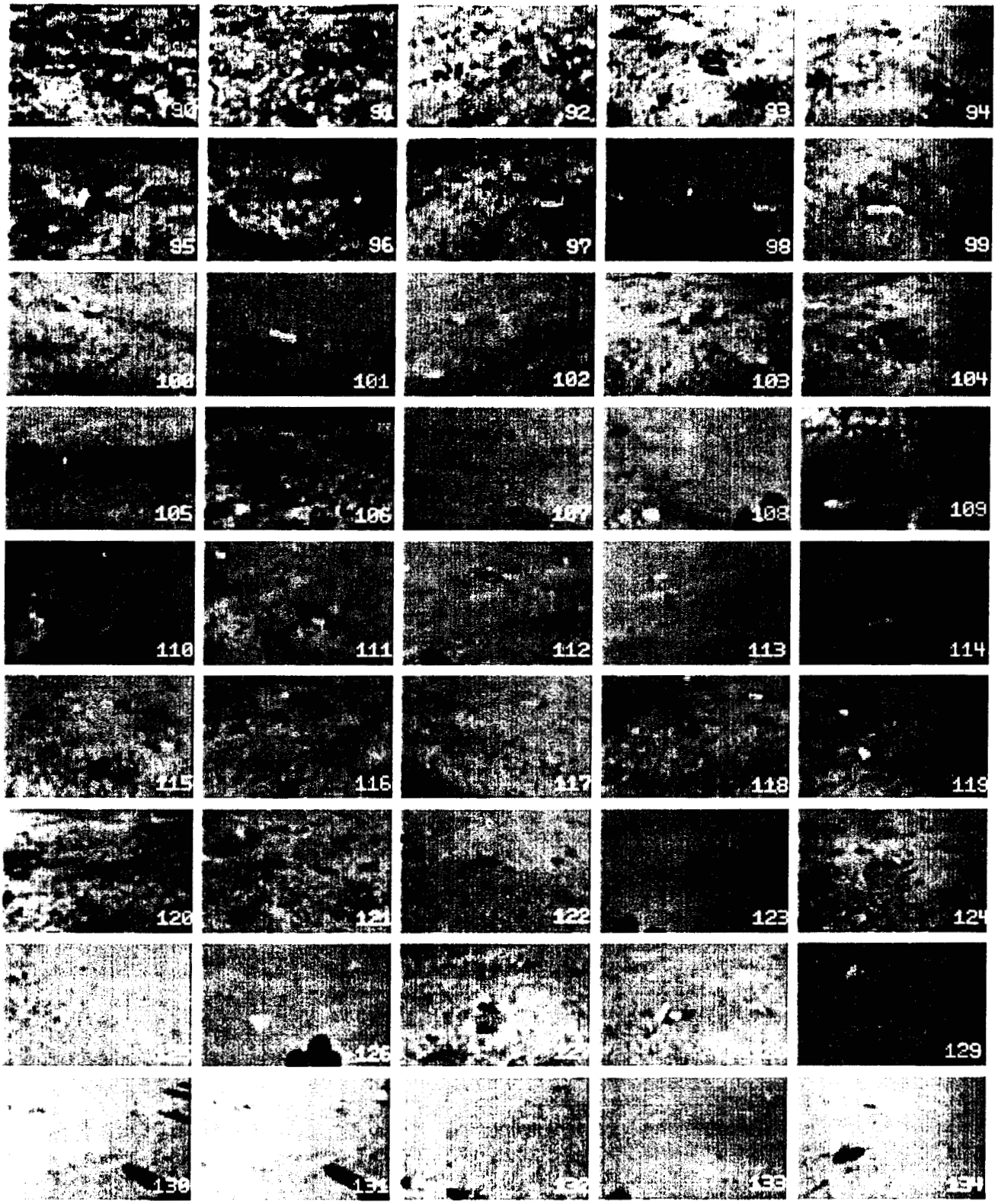

Figure 5: Data set thumbnails 90-134.

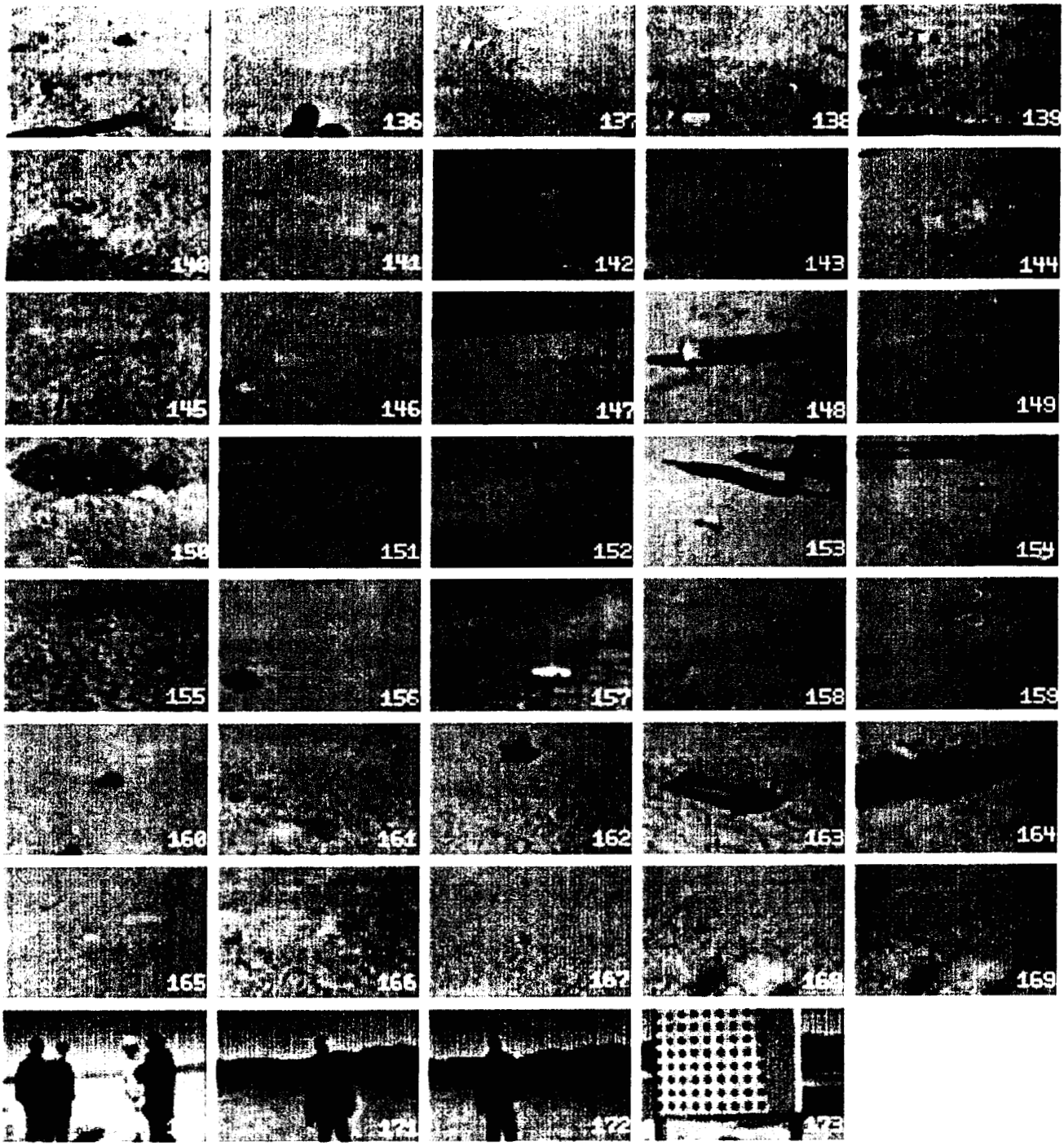

Figure 6: Data set thumbnails 134-173.

Im.	Object	Location	Notes	Im.	Object	Location	Notes
0	calibration target		X	35	BLU-97	(34,440)	D,O,S,X
1	BLU-97	(345,428)		36	BLU-97	(208,154)	D,S,X
2	BLU-97	(539,170)		37	BLU-97	(189.274)	D
3	BLU-97	(427,120)		38	BLU-97	(114,203)	D
4	BLU-97	(496,205)		39	BLU-97	(280,99)	D,T
5	BLU-97	(176,164)		40	BLU-97	(312,418)	D,S
6	BLU-97	(121,405)	S,X	41	BLU-97	(409,238)	D,S
7	BLU-97	(227,284)	S,X	42	BLU-97	(459,164)	D,S
8	BLU-97	(326,75)	S,X	43	BLU-97	(224,383)	D,S
9	BLU-97	(367,356)	C	44	BLU-97	(533,302)	T
10	BLU-97	(615,119)	C	45	BLU-97	(158,380)	
11	BLU-97	(393,324)	C,T	46	BLU-97	(475,183)	S,X
12	BLU-38	(293,412)		47	BLU-97	(244,135)	S,X
13	BLU-97	(472,403)	D,S,X	48	BLU-97	(505,381)	S,X
14	BLU-97	(318,406)	D,S,X	49	BLU-97	(195,129)	S,X
15	BLU-97	(347,201)	D	50	BLU-97	(318,272)	D
16	BLU-97	(245,98)	D	51	BLU-97	(504,446)	D
17	BLU-97	(341,85)	C	52	BLU-97	(322,225)	D
18	BLU-97	(134,125)	C	53	BLU-97	(415,75)	D
19	BLU-97	(426,353)	C,T	54	BLU-97	(270,268)	D,S,X
20	BLU-97	(94,440)	C,S,X	55	BLU-97	(533,180)	D,S,X
21	BLU-97	(219,104)	C,S	56	BLU-97	(218,421)	D,S,X
22	BLU-97	(281,26)	C,S,X	57	no ordnance		
23	BLU-38	(366,370)		58	BLU-97	(118,350)	D,S,X
24	BLU-97	(301,397)	D,S,X	59	BLU-97	(155,225)	D,S,X
25	BLU-97	(368,414)	D,S,X	60	BLU-97	(264,179)	D
26	BLU-97	(616,199)	D,S,X	61	BLU-97	(485,457)	D
27	BLU-97	(353,353)	D,S,X	62	BLU-97	(35,297)	O
28	BLU-97	(54,267)	D,S	63	BLU-97	(330,130)	D,S
29	BLU-97	(245,416)	D,S,X	64	BLU-97	(538,190)	D,S,X
30	BLU-97	(373,285)	D,S,X	65	BLU-97	(369,216)	S,X
31	BLU-97	(328,155)	D,S,X	66	BLU-97	(387,364)	S,X
32	BLU-97	(336,152)	D	67	BLU-97	(328,168)	
33	BLU-97	(244,78)	D	68	BLU-97	(508,275)	S,X
34	BLU-97	(152,173)	D,S,X	69	BLU-97	(396,364)	S

Table 1: Data set annotation (part 1). Key: S - Partially or completely in shadow (includes self-shadowing). C - Discolored (includes dusty). D - Deformed (incomplete, misshapen, badly dented). O - Obscured (partially buried, underwater, occluded, out of image). M - Misfire (stereo images not at same camera location). X - Excluded from culled data set. T - Part of training set.

Im.	Object	Location	Notes	Im.	Object	Location	Notes
70	BLU-97	(125,203)	S		BLU-97	(474,241)	S,X
71	BLU-97	(271,162)	S		other scraps		
72	BLU-97	(400,86)	S,X	92	BLU-97	(275,286)	D
73	BLU-97	(415,287)	S,X		BLU-97	(131,155)	D
74	BLU-97	(358,395)	S,O		BLU-97	(401,60)	O
75	BLU-97	(282,211)	T		BLU-97	(486,87)	D
76	BLU-97	(471,181)			BLU-97	(732,184)	D
77	BLU-97	(335,141)	D,S,X		BLU-97	(724,13)	D
78	BLU-97	(274,310)	D,S,X		other scraps		
79	BLU-97	(286,374)	D,S,X	93	BLU-97	(346,269)	C
80	BLU-97	(295,167)	D	94	BLU-97	(393,156)	C
81	BLU-97	(405,125)	D	95	BLU-97	(349,259)	D,X
82	BLU-97	(500,67)		96	BLU-97	(600,256)	
83	BLU-97	(439,207)		97	BLU-97	(573,273)	T
84	BLU-97	(496,227)	S		BLU-97	(213,175)	S,O
85	BLU-97	(306,277)	S,X		BLU-97	(685,184)	D,O
86	BLU-97	(23,264)	S,O,X	98	BLU-97	(49,229)	
87	BLU-97	(539,230)	S		BLU-97	(350,228)	D
88	BLU-97	(385,165)	S		BLU-97	(543,59)	
89	BLU-97	(174,290)	S,O,X	99	BLU-97	(306,293)	D
	BLU-97	(72,256)	S,O,X	100	BLU-97	(221,113)	
	BLU-97	(164,215)	X	101	BLU-97	(302,260)	T
	BLU-97	(160,120)	D,O,X	102	BLU-97	(227,201)	
	BLU-97	(228,53)	D,S,X	103	BLU-97	(330,201)	
	BLU-97	(448,65)	D,O,X	104	BLU-97	(285,296)	D,S,X
	BLU-97	(604,279)	S,O,X	105	BLU-97	(248,226)	O
	other scraps			106	BLU-97	(284,95)	D,S,X
90	BLU-97	(58,325)	D,S,X	107	BLU-38	(422,440)	
	BLU-97	(399,309)	O,X	108	BLU-97	(204,435)	D,C,T
	BLU-97	(692,406)	D,S,X		BLU-97	(567,107)	D,C,T
	BLU-97	(666,125)	D,S,X		BLU-97	(14,58)	O,C,T
	BLU-97	(483,76)	D,S,X	109	BLU-97	(302,408)	O,X
	other scraps			110	BLU-97	(290,412)	O,X
91	BLU-97	(150,109)	D,S,X		BLU-97	(408,73)	O,S,X
	BLU-97	(330,77)	D,X	111	BLU-97	(460,330)	D,C
	BLU-97	(454,276)	D,X		BLU-97	(84,73)	

Table 2: Data set annotation (part 2). Key: S - Partially or completely in shadow (includes self-shadowing). C - Discolored (includes dusty). D - Deformed (incomplete, misshapen, badly dented). O - Obscured (partially buried, underwater, occluded, out of image). M - Misfire (stereo images not at same camera location). X - Excluded from culled data set. T - Part of training set.

Im.	Object	Location	Notes	Im.	Object	Location	Notes
112	BLU-97	(21,121)	O	140	BLU-97	(268,220)	D,S
	BLU-97	(647,95)	D	141	BLU-97	(171,174)	D,S
113	BLU-97	(230,163)		142	BLU-97	(276,167)	D
114	BLU-97	(313,334)		143	BLU-97	(360,72)	D
115	BLU-97	(413,126)		144	BLU-97	(279,285)	D,T
116	BLU-97	(332,89)		145	BLU-97	(358,225)	D
117	BLU-97	(62,94)		146	BLU-97	(63,328)	D
	BLU-97	(507,126)			BLU-97	(469,21)	D
118	BLU-97	(353,26)		147	no ordnance		M,X
	BLU-97	(620,98)		148	no ordnance		
119	BLU-97	(192,159)		149	BLU-97	(433,191)	D,S
	BLU-97	(266,323)	O	150	no ordnance		
120	BLU-97	(17,351)	O,S,X	151	no ordnance		
	BLU-97	(556,183)	S,X	152	BLU-97	(544,122)	D
121	BLU-97	(164,103)		153	BLU-97	(277,345)	D
	BLU-97	(489,254)	D	154	BLU-97	(506,209)	D
122	BLU-97	(616,217)	D,S,X	155	BLU-97	(199,189)	D
123	BLU-97	(712.154)	D,S,X	156	no ordnance		
124	BLU-97	(366,229)	S	157	no ordnance		
125	BLU-97	(57,150)	S	158	no ordnance		
126	BLU-97	(623,114)		159	no ordnance		
127	BLU-97	(483,95)		160	no ordnance		
128	BLU-97	(316,305)	D	161	BLU-97	(85,290)	D,S
129	BLU-97	(268,134)	D	162	no ordnance		
130	BLU-97	(88,130)		163	no ordnance		
131	BLU-97	(112,112)	M,X	164	no ordnance		
132	BLU-97	(130,439)		165	no ordnance		
	BLU-97	(340,30)	D	166	BLU-97	(111,271)	O,S,X
133	BLU-97	(683,91)	D	167	BLU-97	(339,281)	M,X
134	BLU-97	(584,297)	D,T	168	BLU-97	(232,239)	M,X
135	BLU-97	(146,309)	D,S	169	BLU-97	(230,240)	D,S,X
	BLU-97	(364,78)	D	170	no ordnance		
136	BLU-97	(136,54)	D	171	no ordnance		
137	BLU-97	(185,155)	D	172	no ordnance		
138	BLU-97	(199,427)	D	173	calibration target		X
139	BLU-97	(254,138)	D				

Table 3: Data set annotation (part 3). Key: S - Partially or completely in shadow (includes self-shadowing). C - Discolored (includes dusty). D - Deformed (incomplete, misshapen, badly dented). O - Obscured (partially buried, underwater, occluded, out of image). M - Misfire (stereo images not at same camera location). X - Excluded from culled data set. T - Part of training set.

$$y = \frac{(\mathbf{p} - \mathbf{c}) \cdot \mathbf{v}}{(\mathbf{p} - \mathbf{c}) \cdot \mathbf{a}}$$

When radial lens distortion is included, the following 3-vectors are added to the model:

- **o**: The optical axis. (If the image sensor plane is perpendicular to the optical axis, then $\mathbf{o} = \mathbf{a}$.)
- ρ: The distortion coefficients. (Coefficients above the second order are neglected.)

Now, the image coordinates of a point $\mathbf{p}$ in free space are given by:

$$\zeta = (\mathbf{p} - \mathbf{c}) \cdot \mathbf{o}$$
$$\lambda = \mathbf{p} - \mathbf{c} - \zeta\mathbf{o}$$
$$\tau = \frac{\lambda \cdot \lambda}{(\mathbf{p}' - \mathbf{c}) \cdot \mathbf{o}}$$
$$\mu \approx \rho_0 + \rho_1\tau + \rho_2\tau^2$$
$$\mathbf{p} = \mathbf{p} + \mu\lambda$$
$$x = \frac{(\mathbf{p}' - \mathbf{c}) \cdot \mathbf{h}}{(\mathbf{p}' - \mathbf{c}) \cdot \mathbf{a}}$$
$$y = \frac{(\mathbf{p}' - \mathbf{c}) \cdot \mathbf{v}}{(\mathbf{p}' - \mathbf{c}) \cdot \mathbf{a}}$$

See [1] for further details.
For this dataset, camera calibration yielded the following values for these variables.

Left camera

$\mathbf{c}$ = {	-0.120648	-0.001099	1.293646 }
$\mathbf{a}$ = {	0.000812	0.966538	-0.256522 }
$\mathbf{h}$ = {	1585.295146	389.091457	-96.430581 }
$\mathbf{v}$ = {	6.941665	-174.786624	-1593.129471 }
$\mathbf{o}$ = {	0.001827	0.965911	-0.258868 }
ρ = {	0.000005	-0.126333	-0.049222 }

Right camera

$\mathbf{c}$ = {	0.120648	0.001099	1.295894 }
$\mathbf{a}$ = {	-0.000812	0.965308	-0.261114 }
$\mathbf{h}$ = {	1589.006034	375.532685	-93.899922 }
$\mathbf{v}$ = {	7.802359	-175.325657	-1598.488874 }
$\mathbf{o}$ = {	-0.003243	0.963327	-0.268308 }
ρ = {	0.000023	-0.160986	0.907825 }

Note that these camera models point the cameras in the positive y-direction, with a downward tilt of 15 degrees. The x-axis is to the right and the z-axis is up. The cameras stand approximately 1.3 meters above the ground on the tripod.

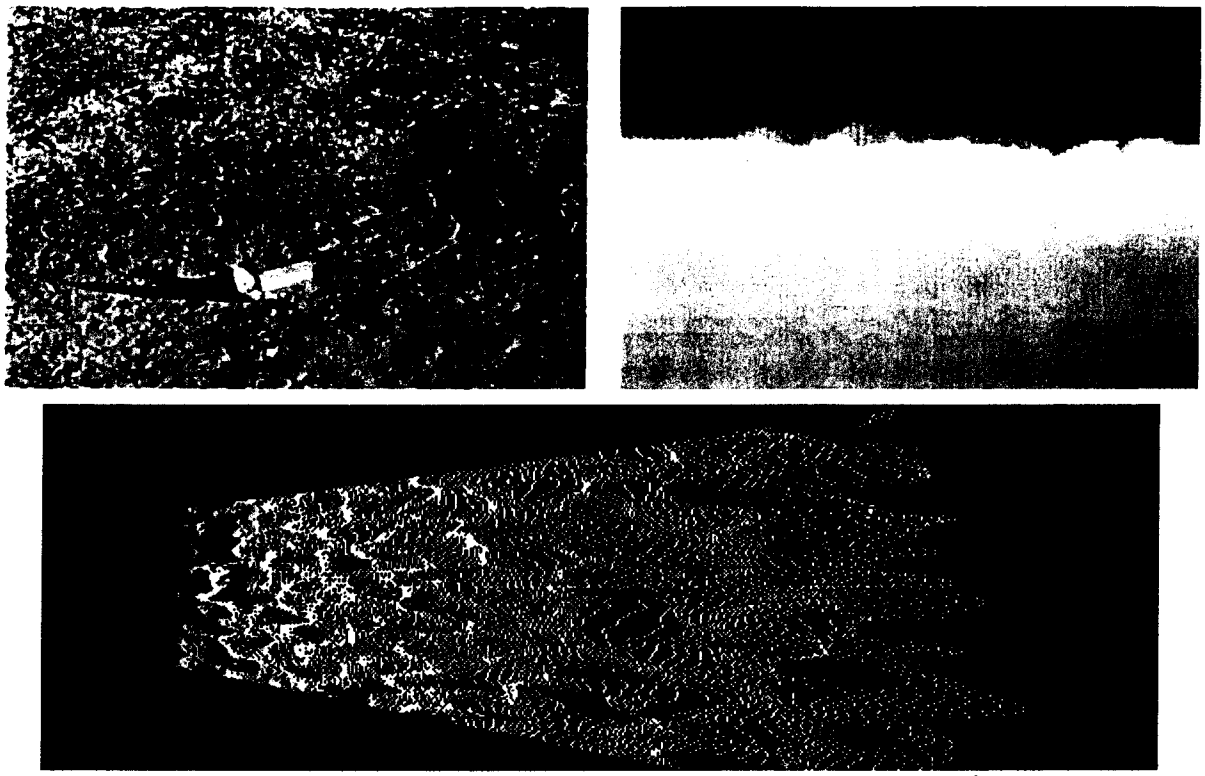

Figure 7: Range data extracted from image 9. (a) The left image of the stereo pair. (b) False color map of the distance to each pixel. Note that pixels without range estimates have been filled by a process that gives precedence to horizontal directions over vertical directions. (c) Overhead plot of the range points extracted from the stereo pair. The color is an indication of the height of the point.

4.2 Range data

Given the camera parameters, we are able to extract range data from the stereo pairs using an algorithm developed by Larry Matthies and implemented by Todd Litwin [3, 4]. Each image is first warped to remove the lens distortion and the images are rectified so that the corresponding scan-lines yield corresponding epipolar lines in the image. The disparity between the left and right images is measured for each pixel by minimizing the sum-of-squared-difference (SSD) measure of windows around the pixel in the Laplacian of the image over a finite disparity range. Subpixel disparity estimates are computed using parabolic interpolation on the SSD values neighboring the minimum. Outliers are removed through consistency checking and smoothing is performed over a 3×3 window to reduce noise. Finally, the coordinates of each pixel are computed using triangulation. Figure 6 shows an example of the range data extracted for a stereo pair in this data set.

5 Summary

This report has described a data set that was collected for the purpose of evaluating ordnance recognition algorithms. The data set consists of 174 color stereo images of (primarily) BLU-97 ordnance at a test range near Nellis Air Force Base.

Acknowledgments

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by the Wright Laboratory at Tyndall Air Force Base, Panama City, Florida, through an agreement with the National Aeronautics and Space Administration.

References

[1] D. B. Gennery. Least-squares camera calibration including lens distortion and automatic editing of calibration points. In A. Grün and T. S. Huang, editors, *Calibration and Orientation of Cameras in Computer Vision*. Springer-Verlag, in press.

[2] C. M. Mackenzie, C. Jordan, R. E. Dugan, and M. A. Kolodny. Detecting UXO - putting it all into perspective. In *Detection Technologies for Mines and Minelike Targets, Proc. SPIE 2496*, pages 94–99, 1995.

[3] L. Matthies. Stereo vision for planetary rovers: Stochastic modeling to near real-time implementation. *International Journal of Computer Vision*, 8(1):71–91, July 1992.

[4] L. Matthies, A. Kelly, T. Litwin, and G. Tharp. Obstacle detection for unmanned ground vehicles: A progress report. In *Proceedings of the International Symposium on Robotics Research*, pages 475–486, 1996.

[5] C. F. Olson and Larry H. Matthies. Visual ordnance recognition for clearing test ranges. In *Detection and Remediation Technologies for Mines and Mine-Like Targets III, Proc. SPIE*, 1998.

[6] Y. Yakimovsky and R. Cunningham. A system for extracting three-dimensional measurements from a stereo pair of TV cameras. *Computer Vision, Graphics, and Image Processing*, 7:195–210, 1978.